SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Graham, Laurie A.
Liou, Yih-Cherng
Walker, Virginia K.
Davies, Peter L.

(ii) TITLE OF INVENTION: Tenebrio Antifreeze Proteins

(iii) NUMBER OF SEQUENCES: 22

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Townsend and Townsend and Crew LLP
(B) STREET: Two Embarcadero Center, Eighth Floor
(C) CITY: San Francisco
(D) STATE: California
(E) COUNTRY: USA
(F) ZIP: 94111-3834

A10

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/882,907
(B) FILING DATE: 26-JUN-1997
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Weber, Kenneth A.
(B) REGISTRATION NUMBER: 31,677
(C) REFERENCE/DOCKET NUMBER: 016252-002100US

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (415) 576-0200
(B) TELEFAX: (415) 576-0300

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Peptide
(B) LOCATION: 1..12
(D) OTHER INFORMATION: /note= "consensus 12 amino acid repeating motif"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Cys Thr Xaa Ser Xaa Xaa Cys Xaa Xaa Ala Xaa Thr
1 5 10

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

A,10

(ix) FEATURE:
(A) NAME/KEY: -
(B) LOCATION: 1..68
(D) OTHER INFORMATION: /note= "conserved N-terminal region of thermal hysteresis protein (THP) isoforms"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TGCACTGGGG STGCTGATTG YACTAGTTGT ACAGVWGCAT GCACTGGTTG TGGAARYTGT 60

CCAAATGC 68

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Peptide
(B) LOCATION: 1..16
(D) OTHER INFORMATION: /note= "consensus 16 amino acid N-terminal motif for YL-1, YL-2, YL-3 and YL-4"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Xaa Cys Thr Xaa Xaa Xaa Xaa Cys Thr Xaa Cys Thr Xaa Xaa Cys Thr
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:
(A) NAME/KEY: Peptide
(B) LOCATION: 1..24
(D) OTHER INFORMATION: /note= "N-terminal amino acid sequence of YL-1, YL-2, YL-3 and YL-4"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 1
(D) OTHER INFORMATION: /product= "OTHER"
/note= "Xaa = Gln or His"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /product= "OTHER"
/note= "Xaa = Ala or Gly"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 13
(D) OTHER INFORMATION: /product= "OTHER"
/note= "Xaa = Ala, Asp or Gly"

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 20
(D) OTHER INFORMATION: /product= "OTHER"
/note= "Xaa = Asn or Ser"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Xaa Cys Thr Gly Xaa Ala Asp Cys Thr Ser Cys Thr Xaa Ala Cys Thr
1               5                  10                  15

Gly Cys Gly Xaa Cys Pro Asn Ala
            20
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 61 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: -
(B) LOCATION: 1..61
(D) OTHER INFORMATION: /note= "conserved C-terminal region of thermal hysteresis protein (THP) isoforms"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CTCAACCAAC TGTTACAAAG CTACAGCCTG TACCAATTCA WCAGGATGTC CCGGACATTA 60

R 61

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: -
(B) LOCATION: 1..32
(D) OTHER INFORMATION: /note= "PCR primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CATATGCATA TGCAATGCAC TGGGGGTGCT GA 32

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: -
(B) LOCATION: 1..32
(D) OTHER INFORMATION: /note= "PCR primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AAGCCTAAGC TTTTAATGTC CGGGACATCC TG 32

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: -
(B) LOCATION: 1..18
(D) OTHER INFORMATION: /note= "internal downstream facing sequencing primer"

A,10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AAAGACTGTT TTGAAGCC 18

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: -
(B) LOCATION: 1..18
(D) OTHER INFORMATION: /note= "internal upstream facing sequencing primer"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TTCAAAACAG TCTTTTGA 18

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 559 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: -
(B) LOCATION: 1..559
(D) OTHER INFORMATION: /note= "YL-1 thermal hysteresis protein (THP) cDNA from Tenebrio molitor"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 37..375
(D) OTHER INFORMATION: /product= "YL-1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
TAAACAGCGA GATAAACAAC AATACTACAT AAAACT ATG GCG TTC AAA ACG TGT        54
                                          Met Ala Phe Lys Thr Cys
                                            1               5

GGT TTT TCA AAA AAA TGG TTA GTA ATA GCA GTT ATA GTT ATG TGT TTG       102
Gly Phe Ser Lys Lys Trp Leu Val Ile Ala Val Ile Val Met Cys Leu
            10                  15                  20

TGT ACC GAG TGT TAT TGC CAC TGC ACT GGG GGT GCT GAT TGT ACT AGT       150
Cys Thr Glu Cys Tyr Cys His Cys Thr Gly Gly Ala Asp Cys Thr Ser
        25                  30                  35

TGT ACA GAT GCA TGC ACT GGT TGT GGA AAT TGT CCA AAT GCA CAT ACG       198
Cys Thr Asp Ala Cys Thr Gly Cys Gly Asn Cys Pro Asn Ala His Thr
    40                  45                  50

TGT ACC GAT TCC AAA AAT TGT GTC AAG GCA GCA ACA TGT ACT GGA TCT       246
Cys Thr Asp Ser Lys Asn Cys Val Lys Ala Ala Thr Cys Thr Gly Ser
55                  60                  65                  70

ACA AAA TGT AAT ACC GCC AGG ACG TGT ACA AAC TCA AAA GAC TGT TTT       294
Thr Lys Cys Asn Thr Ala Arg Thr Cys Thr Asn Ser Lys Asp Cys Phe
                75                  80                  85

GAA GCC AAA ACA TGT ACT GAC TCA ACC AAC TGT TAC AAA GCT ACA GCC       342
Glu Ala Lys Thr Cys Thr Asp Ser Thr Asn Cys Tyr Lys Ala Thr Ala
            90                  95                  100

TGT ACC AAT TCA ACA GGA TGT CCC GGA CAT TAAGTTTTTC TATTGTCAAC         392
Cys Thr Asn Ser Thr Gly Cys Pro Gly His
        105                 110

AATAATAAAA CACACTTACT GTTATCTTAG CTAAAACATA ATTGTAAGCT CACTCTGTTT     452

TGTATCCTAT CTGTCTCTGC CTCCGAAGGA TGATAATTTT GTACTGGGAG CGAAAGGTTT     512

ATCCGACAAT AATAAACTAA AATAATTGAT ATAAAAAAAA AAAAAAA                   559
```

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 112 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
Met Ala Phe Lys Thr Cys Gly Phe Ser Lys Lys Trp Leu Val Ile Ala
 1               5                  10                  15

Val Ile Val Met Cys Leu Cys Thr Glu Cys Tyr Cys His Cys Thr Gly
            20                  25                  30

Gly Ala Asp Cys Thr Ser Cys Thr Asp Ala Cys Thr Gly Cys Gly Asn
        35                  40                  45

Cys Pro Asn Ala His Thr Cys Thr Asp Ser Lys Asn Cys Val Lys Ala
    50                  55                  60

Ala Thr Cys Thr Gly Ser Thr Lys Cys Asn Thr Ala Arg Thr Cys Thr
65                  70                  75                  80

Asn Ser Lys Asp Cys Phe Glu Ala Lys Thr Cys Thr Asp Ser Thr Asn
                85                  90                  95

Cys Tyr Lys Ala Thr Ala Cys Thr Asn Ser Thr Gly Cys Pro Gly His
            100                 105                 110
```

A,10

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 529 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: -
(B) LOCATION: 1..529
(D) OTHER INFORMATION: /note= "YL-2 thermal hysteresis protein (THP) cDNA from Tenebrio molitor"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 21..359
(D) OTHER INFORMATION: /product= "YL-2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
AAACAATATT ACAAAAAACT ATG GCA TTC AAA ACG TGT GGT TTT TCA AAA          50
                      Met Ala Phe Lys Thr Cys Gly Phe Ser Lys
                        1               5                  10

AAA TGG TTA GTA ATA GCA GTT ATA GTT ATG TGT TTG TGT ACC GAG TGT        98
Lys Trp Leu Val Ile Ala Val Ile Val Met Cys Leu Cys Thr Glu Cys
                15                  20                  25

TAT TGC CAA TGC ACT GGG GGT GCT GAT TGC ACT AGT TGT ACA GGA GCA       146
Tyr Cys Gln Cys Thr Gly Gly Ala Asp Cys Thr Ser Cys Thr Gly Ala
            30                  35                  40
```

```
TGC ACT GGT TGT GGA AAC TGT CCA AAT GCA GTA ACG TGT ACC AAT TCT        194
Cys Thr Gly Cys Gly Asn Cys Pro Asn Ala Val Thr Cys Thr Asn Ser
          45                  50                  55

CAA CAT TGT GTC AAG GCA AAT ACA TGT ACT GGG TCT ACA GAT TGT AAT        242
Gln His Cys Val Lys Ala Asn Thr Cys Thr Gly Ser Thr Asp Cys Asn
     60                  65                  70

ACA GCC CAG ACG TGT ACA AAC TCA AAA GAC TGT TTT GAA GCC AAC ACA        290
Thr Ala Gln Thr Cys Thr Asn Ser Lys Asp Cys Phe Glu Ala Asn Thr
 75                  80                  85                  90

TGT ACT GAC TCA ACC AAC TGT TAC AAA GCT ACA GCC TGT ACC AAT TCA        338
Cys Thr Asp Ser Thr Asn Cys Tyr Lys Ala Thr Ala Cys Thr Asn Ser
                 95                 100                 105

TCA GGA TGT CCC GGA CAT TAAGTTTTTC TATTGTCAAC AATCATAAAA               386
Ser Gly Cys Pro Gly His
            110

CACAATTATT GTTAGCTAAG TTAAAACTCT GTATTGTATC CGATCTGTCT CTTTGCCTCC      446

CAAGGATGAT AATTTTGTAC TGGGAGCGAA AGGGTTATCG GACAATAATA AACTAAAATA      506

ATTGATATAA AAAAAAAAAA AAA                                              529
```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 112 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
Met Ala Phe Lys Thr Cys Gly Phe Ser Lys Lys Trp Leu Val Ile Ala
 1               5                  10                  15

Val Ile Val Met Cys Leu Cys Thr Glu Cys Tyr Cys Gln Cys Thr Gly
            20                  25                  30

Gly Ala Asp Cys Thr Ser Cys Thr Gly Ala Cys Thr Gly Cys Gly Asn
        35                  40                  45

Cys Pro Asn Ala Val Thr Cys Thr Asn Ser Gln His Cys Val Lys Ala
    50                  55                  60

Asn Thr Cys Thr Gly Ser Thr Asp Cys Asn Thr Ala Gln Thr Cys Thr
65                  70                  75                  80

Asn Ser Lys Asp Cys Phe Glu Ala Asn Thr Cys Thr Asp Ser Thr Asn
                85                  90                  95

Cys Tyr Lys Ala Thr Ala Cys Thr Asn Ser Ser Gly Cys Pro Gly His
            100                 105                 110
```

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 642 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: -
(B) LOCATION: 1..642
(D) OTHER INFORMATION: /note= "YL-4 thermal hysteresis protein (THP) cDNA from Tenebrio molitor"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 10..456
(D) OTHER INFORMATION: /product= "YL-4"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
CAAAAAAGT ATG TCA TTC AAA ATA AGT ACT TTT ACA AAA ATC TGG TTA          48
          Met Ser Phe Lys Ile Ser Thr Phe Thr Lys Ile Trp Leu
            1               5                  10

ATT ATA GCA GTT ATC GTT ATG TGT TTG TGT AAC GAG TAT AAT TGC CAG        96
Ile Ile Ala Val Ile Val Met Cys Leu Cys Asn Glu Tyr Asn Cys Gln
     15                  20                  25

TGC ACT GGG GCT GCT GAT TGT ACT AGT TGT ACA GCA GCA TGC ACT GGT       144
Cys Thr Gly Ala Ala Asp Cys Thr Ser Cys Thr Ala Ala Cys Thr Gly
 30                  35                  40                  45

TGT GGA AAC TGT CCA AAT GCA ATA ACG TGT ACC GGT TCT AAA AAT TGT       192
Cys Gly Asn Cys Pro Asn Ala Ile Thr Cys Thr Gly Ser Lys Asn Cys
                 50                  55                  60

GTC AGG GCA ACA ACA TGT ACT GGG TCT ACA AAC TGT AAT AGA GCC ACG       240
Val Arg Ala Thr Thr Cys Thr Gly Ser Thr Asn Cys Asn Arg Ala Thr
            65                  70                  75

ACG TGT ACA AAT TCA AAA GGC TGT TTA GAA GCC ACA ACA TGT ACT GGG       288
Thr Cys Thr Asn Ser Lys Gly Cys Leu Glu Ala Thr Thr Cys Thr Gly
         80                  85                  90

TCT ACA CAC TGT CAT AGA GCC ACG ACG TGT ACA AAT TCA AAA GAC TGT       336
Ser Thr His Cys His Arg Ala Thr Thr Cys Thr Asn Ser Lys Asp Cys
     95                  100                 105

TTT GAA GCC ACA ACA TGT ACT GGC TCA AGC AAC TGT TAC ACT GCT ACA       384
Phe Glu Ala Thr Thr Cys Thr Gly Ser Ser Asn Cys Tyr Thr Ala Thr
110                 115                 120                 125

ACA TGT ACT AAC TCA ACC AAC TGT TAC AAA GCT ACA GCC TGT ACC AAT       432
Thr Cys Thr Asn Ser Thr Asn Cys Tyr Lys Ala Thr Ala Cys Thr Asn
                130                 135                 140
```

TCA ACA GGA TGT CCC GGA CAT TAGGTTTTTT TATTGTCAAC AATAAAATAA 483
Ser Thr Gly Cys Pro Gly His
145

AACAAAACTG TTCTTATCTA AGCTAAAACA TAAATGTAAA CGTTAATTTG TATTCTATCC 543

GATCTGTCCC TTTGCGCCCT AAGGATAATT TTGTACAGGG AGAGAAAAGG CTATCGGACA 603

ATAATAAACA TTGTTAATAT ACATAAAAAA AAAAAAAAA 642

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 148 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

Q10

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Met Ser Phe Lys Ile Ser Thr Phe Thr Lys Ile Trp Leu Ile Ile Ala
1 5 10 15

Val Ile Val Met Cys Leu Cys Asn Glu Tyr Asn Cys Gln Cys Thr Gly
20 25 30

Ala Ala Asp Cys Thr Ser Cys Thr Ala Ala Cys Thr Gly Cys Gly Asn
35 40 45

Cys Pro Asn Ala Ile Thr Cys Thr Gly Ser Lys Asn Cys Val Arg Ala
50 55 60

Thr Thr Cys Thr Gly Ser Thr Asn Cys Asn Arg Ala Thr Thr Cys Thr
65 70 75 80

Asn Ser Lys Gly Cys Leu Glu Ala Thr Thr Cys Thr Gly Ser Thr His
85 90 95

Cys His Arg Ala Thr Thr Cys Thr Asn Ser Lys Asp Cys Phe Glu Ala
100 105 110

Thr Thr Cys Thr Gly Ser Ser Asn Cys Tyr Thr Ala Thr Thr Cys Thr
115 120 125

Asn Ser Thr Asn Cys Tyr Lys Ala Thr Ala Cys Thr Asn Ser Thr Gly
130 135 140

Cys Pro Gly His
145

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 587 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: -
(B) LOCATION: 1..587
(D) OTHER INFORMATION: /note= "YL-3 thermal hysteresis protein (THP) cDNA from Tenebrio molitor"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 24..398
(D) OTHER INFORMATION: /product= "YL-3"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```
AAACAACAAT ATTACAAAAA ACT ATG GCA TTC AAA ACG TGT GGT TTT TCA          50
                          Met Ala Phe Lys Thr Cys Gly Phe Ser
                            1               5

AAA AAA TGG TTA ATA ATA GCA GTT ATA GTT ATG TGT TTG TGT ACC GAG        98
Lys Lys Trp Leu Ile Ile Ala Val Ile Val Met Cys Leu Cys Thr Glu
 10                  15                  20                  25

TGT TAT TGC CAA TGC ACT GGG GGT GCT GAT TGT ACT AGT TGT ACA GCA       146
Cys Tyr Cys Gln Cys Thr Gly Gly Ala Asp Cys Thr Ser Cys Thr Ala
                 30                  35                  40

GCA TGC ACT GGT TGT GGA AGT TGT CCA AAT GCG CAT ACG TGT ACC GAT       194
Ala Cys Thr Gly Cys Gly Ser Cys Pro Asn Ala His Thr Cys Thr Asp
             45                  50                  55

TCT AAA AAT TGT GTC AGG GCA GAA ACG TGT ACC GAT TCT GAA AAT TGT       242
Ser Lys Asn Cys Val Arg Ala Glu Thr Cys Thr Asp Ser Glu Asn Cys
         60                  65                  70

GTC AAG GCA CAT ACA TGT ACT GGA TCT AGA AAC TGT AAT ACA GCC ATG       290
Val Lys Ala His Thr Cys Thr Gly Ser Arg Asn Cys Asn Thr Ala Met
     75                  80                  85

ACG TGT ACA AAC TCA AAA GAC TGT TTT GAA GCC AAA ACA TGT ACT GAC       338
Thr Cys Thr Asn Ser Lys Asp Cys Phe Glu Ala Lys Thr Cys Thr Asp
 90                  95                 100                 105

TCA ACC AAC TGT TAC AAA GCT ACA GCC TGT ACC AAT TCA ACA GGA TGT       386
Ser Thr Asn Cys Tyr Lys Ala Thr Ala Cys Thr Asn Ser Thr Gly Cys
                110                 115                 120

CCC GGA CAT TAAGTTTTTC TATTGTCAAC AATAATAAAA CACGGAGGGA               435
Pro Gly His

TAGTCTAAGC TAAAACATAA TTGTAAGCTT ACTCTGTATT GTATCCGATC TGTCTCTTTG     495

CCTCCCAAGG ATGATAATTT TGTACTGGGA GCGAAAGGGT TACCGGACAA TAATAATTAA     555

TAAACTAAAT AATTGATAAA AAAAAAAAAA AA                                   587
```

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 124 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```
Met Ala Phe Lys Thr Cys Gly Phe Ser Lys Lys Trp Leu Ile Ile Ala
 1               5                  10                  15

Val Ile Val Met Cys Leu Cys Thr Glu Cys Tyr Cys Gln Cys Thr Gly
            20                  25                  30

Gly Ala Asp Cys Thr Ser Cys Thr Ala Ala Cys Thr Gly Cys Gly Ser
        35                  40                  45

Cys Pro Asn Ala His Thr Cys Thr Asp Ser Lys Asn Cys Val Arg Ala
    50                  55                  60

Glu Thr Cys Thr Asp Ser Glu Asn Cys Val Lys Ala His Thr Cys Thr
65                  70                  75                  80

Gly Ser Arg Asn Cys Asn Thr Ala Met Thr Cys Thr Asn Ser Lys Asp
                85                  90                  95

Cys Phe Glu Ala Lys Thr Cys Thr Asp Ser Thr Asn Cys Tyr Lys Ala
            100                 105                 110

Thr Ala Cys Thr Asn Ser Thr Gly Cys Pro Gly His
        115                 120
```

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 602 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: -
(B) LOCATION: 1..602
(D) OTHER INFORMATION: /note= "5-15 cDNA"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 36..410
(D) OTHER INFORMATION: /product= "5-15"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```
AAACAGCGAG ATAAACAACA ATATTACAAA AAACT ATG GCA TTC AAA ACG TGT          53
                                       Met Ala Phe Lys Thr Cys
                                         1               5

GGT TTT TCA AAA AAA TGG TTA ATA ATA GCA GTT ATA GTT ATG TGT TTG        101
Gly Phe Ser Lys Lys Trp Leu Ile Ile Ala Val Ile Val Met Cys Leu
            10                  15                  20

TGT ACC GAG TGT TAT TGC CAA TGC ACT GGG GGT GCT GAT TGT ACT AGT        149
Cys Thr Glu Cys Tyr Cys Gln Cys Thr Gly Gly Ala Asp Cys Thr Ser
        25                  30                  35

TGT ACA GCA GCA TGC ACT GGT TGT GGA AGT TGT CCA AAT GCG CAT ACG        197
Cys Thr Ala Ala Cys Thr Gly Cys Gly Ser Cys Pro Asn Ala His Thr
    40                  45                  50

TGT ATC GAT TCT AAA AAT TGT GTC AGG GCA GAA ACG TGT ACC GAT TCT        245
Cys Ile Asp Ser Lys Asn Cys Val Arg Ala Glu Thr Cys Thr Asp Ser
55                  60                  65                  70

GAA AAT TGT GTC AAG GCA CAT ACA TGT ACT GGA TCT AGA AAC TGT AAT        293
Glu Asn Cys Val Lys Ala His Thr Cys Thr Gly Ser Arg Asn Cys Asn
                75                  80                  85

ACA GCC ATG ACG TGT ACA AAC TCA AAA GAC TGT TTT GAA GCC AAA ACA        341
Thr Ala Met Thr Cys Thr Asn Ser Lys Asp Cys Phe Glu Ala Lys Thr
            90                  95                  100

TGT ACT GAC TCA ACC AAC TGT TAC AAA GCT ACA GCC TGT ACC AAT TCA        389
Cys Thr Asp Ser Thr Asn Cys Tyr Lys Ala Thr Ala Cys Thr Asn Ser
        105                 110                 115

ACA GGA TGT CCC GGA CAT TAAGTTTTTC TATTGTCAAC AATAATAAAA               437
Thr Gly Cys Pro Gly His
    120

CACGGAGGGA TAGTCTAAGC TAAAACATAA TTGTAAGCTT ACTCTGTATT GTATCCGATC      497

TGTCTCTTTG CCTCCCAAGG ATGATAATTT TGTACTGGGA GCGAAAGGGT TACCGGACAA      557

TAATAATTAA TAAACTAAAA TAATTGATAT AAAAAAAAAA AAAAA                      602
```

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 124 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```
Met Ala Phe Lys Thr Cys Gly Phe Ser Lys Lys Trp Leu Ile Ile Ala
 1               5                  10                  15

Val Ile Val Met Cys Leu Cys Thr Glu Cys Tyr Cys Gln Cys Thr Gly
            20                  25                  30
```

Gly Ala Asp Cys Thr Ser Cys Thr Ala Ala Cys Thr Gly Cys Gly Ser
35 40 45

Cys Pro Asn Ala His Thr Cys Ile Asp Ser Lys Asn Cys Val Arg Ala
50 55 60

Glu Thr Cys Thr Asp Ser Glu Asn Cys Val Lys Ala His Thr Cys Thr
65 70 75 80

Gly Ser Arg Asn Cys Asn Thr Ala Met Thr Cys Thr Asn Ser Lys Asp
85 90 95

Cys Phe Glu Ala Lys Thr Cys Thr Asp Ser Thr Asn Cys Tyr Lys Ala
100 105 110

Thr Ala Cys Thr Asn Ser Thr Gly Cys Pro Gly His
115 120

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 555 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: -
(B) LOCATION: 1..555
(D) OTHER INFORMATION: /note= "consensus thermal hysteresis protein (THP) cDNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

AAACAGCGAG ATAAANAACA ATANTACANA AAANTATGNC NTTCAAAANN NGTNNTTTTN 60

CAAAAANNTG GTTANTNATA GCAGTTATNG TTATGTGTTT GTGTANCGAG TNTNATTGCC 120

ANTGCACTGG GGNTGCTGAT TGNACTAGTT GTACAGNNGC ATGCACTGGT TGTGGAANNT 180

GTCCAAATGC NNNNACGTGT ANCNNTTCNN AANATTGTGT CANGGCANNN ACATGTACTG 240

GNTCTANANA NTGTAATANN GCCNNGACGT GTACAAANTC AAAAGNCTGT TTNGAAGCCA 300

NNACATGTAC TGACTCAACC AACTGTTACA AAGCTACAGC CTGTACCAAT TCANCAGGAT 360

GTCCCGGACA TTANGTTTTT NTATTGTCAA CAATNNNATA AAACANNNNN NNNNTNNNCT 420

NAGNTAAAAC NNNNNTGTAN NNNNNNNNNT GTNTNNNATC NNNNNTGTCN CTNNGCNNCC 480

NAAGGANNNT AATTTTGTAC NGGGAGNGAA ANGNNTANCN GACAATAATA NNNNNNNAAC 540

NNNNNTAATN NNNAT 555

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 313 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:
(A) NAME/KEY: -
(B) LOCATION: 1..313
(D) OTHER INFORMATION: /note= "Yl-1 nucleic acid probe used to screen Tenebrio molitor larval fat body lambda-Zap cDNA library"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CACTGCACTG GGGGTGCTGA TTGTACTAGT TGTACAGATG CATGCACTGG TTGTGGAAAT 60

TGTCCAAATG CACATACGTG TACCGATTCC AAAAATTGTG TCAAGGCAGC AACATGTACT 120

GGATCTACAA AATGTAATAC CGCCAGGACG TGTACAAACT CAAAAGACTG TTTTGAAGCC 180

AAAACATGTA CTGACTCAAC CAACTGTTAC AAAGCTACAG CCTGTACCAA TTCAACAGGA 240

TGTCCCGGAC ATTAAGTTTT TCTATTGTCA ACAATAATAA AACACACTTA CTGTTATCTT 300

AGCTAAAACA TAA 313

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Cys Xaa Xaa Xaa Xaa Xaa Cys Xaa Xaa Cys Xaa Xaa Xaa Cys Xaa Xaa
1 5 10 15

Cys Xaa Xaa Cys
20

SEQ ID NO:1 12 AMINO ACID MOTIF

Cys-Thr-Xaa-Ser-Xaa-Xaa-Cys-Xaa-Xaa-Ala-Xaa-Thr; where Xaa represents any amino acid

SEQ ID NO:2 CONSERVED N-TERMINAL REGION OF THP ISOFORMS

5'-TGCACTGGGG(GC)TGCTGATTG(TC)ACTAGTTGTACAG(CN)(AT)GCATGCACTGG TTGTGGAA(AG)(TC)TGTCCAAATGC-3'

SEQ ID NO:3 N-TERMINAL MOTIF FOR YL-1, YL-2, YL-3 and YL4

Xaa-Cys-Thr-Xaa-Xaa-Xaa-Xaa-Cys-Thr-Xaa-Cys-Thr-Xaa-Xaa-Cys-Thr

SEQ ID NO:4 N-TERMINAL AMINO ACID SEQUENCE OF YL-1, YL-2, YL-3 AND YL-4

(Gln/His)-Cys-Thr-Gly-(Ala/Gly)-Ala-Asp-Cys-Thr-Ser-Cys-Thr-(Ala/Asp/Gly)-Ala-Cys-Thr-Gly-Cys-Gly-(Asn/Ser)-Cys-Pro-Asn-Ala

SEQ ID NO:5 CONSERVED C-TERMINAL REGION OF THP ISOFORMS

5'-CTCAACCAACTGTTACAAAGCTACAGCCTGTACCAATTCA(TA)CAGGATGTCCCGGA CATTA(AG)-3'

SEQ ID NO:6 5' PCR PRIMER

5'-CATATGCATATGCAATGCACTGGGGGTGCTGA-3'

SEQ ID NO:7 3' PCR PRIMER

5'-AAGCCTAAGCTTTTAATGTCCGGGACATCCTG-3'

SEQ ID NO:8 INTERNAL DOWNSTREAM FACING SEQUENCING PRIMER

5'-AAAGACTGTTTTGAAGCC-3'

SEQ ID NO:9 INTERNAL UPSTREAM FACING SEQUENCING PRIMER

5'-TTCAAAACAGTCTTTTGA-3'

SEQ ID NO:10 NUCLEIC ACID SEQUENCE OF YL-1

TAAACAGCGA GATAAACAAC AATACTACAT AAAACTATGG CGTTCAAAAC
GTGTGGTTTT TCAAAAAAAT GGTTAGTAAT AGCAGTTATA GTTATGTGTT
TGTGTACCGA GTGTTATTGC CACTGCACTG GGGGTGCTGA TTGTACTAGT
TGTACAGATG CATGCACTGG TTGTGGAAAT TGTCCAAATG CACATACGTG
TACCGATTCC AAAAATTGTG TCAAGGCAGC AACATGTACT GGATCTACAA
AATGTAATAC CGCCAGACGT GTACAAACTC AAAAGACTGT TTTGAAGCCA
AAACATGTAC TGACTCAACC AACTGTTACA AAGCTACAGC CTGTACCAAT
TCAACAGGAT GTCCCGGACA TTAAGTTTTT CTATTGTCAA CAATAATAAA
ACACACTTAC TGTTATCTTA GCTAAAACAT AATTGTAAGC TCACTCTGTT
TTGTATCCTA TCTCACTTAC TGTTATCTTA GCTAAAACAT AATTGTAAGC
TCACTCTGTT TTGTATCCTA TCTGTCTCTG CCTCCGAAGG ATGATAATTT
TGTACTGGGA GCGAAAGGTT TATCCGACAA TAATAAACTA AAATAATTGA
TATAAAAAAA AAAAAAAA

C10

SEQ ID NO:11 AMINO ACID SEQUENCE OF YL-1

MAFKTCGFSKKWLVIAVIVMCLCTECYCHCTGGADCTSCTDA CTGCGNCPNAHT
CTDSKNCVKAAT CTGSTKCNTART CTNSKDCFEAKT CTDSTNCYKATA CTNSTGCPGH

SEQ ID NO:12 NUCLEIC ACID SEQUENCE OF YL-2

AAACAATATT ACAAAAAACT ATGGCATTCA AAACGTGTGG TTTTTCAAAA
AAATGGTTAG TAATAGCAGT TATAGTTATG TGTTTGTGTA CCGAGTGTTA
TTGCCAATGC ACTGGGGGTG CTGATTGCAC TAGTTGTACA GGAGCATGCA
CTGGTTGTGG AAACTGTCCA AATGCAGTAA CGTGTACCAA TTCTCAACAT
TGTGTCAAGG CAAATACATG TACTGGGTCT ACAGATTGTA ATACAGCCCA
GACGTGTACA AACTCAAAAG ACTGTTTTGA AGCCAACACA TGTACTGACT
CAACCAACTG TTACAAAGCT ACAGCCTGTA CCAATTCATC AGGATGTCCC
GGACATTAAG TTTTTCTATT GTCAACAATC ATAAAACACA ATTATTGTTA
GCTAAGTTAA AACTCTGTAT TGTATCCGAT CTCAATTATT GTTAGCTAAG
TTAAAACTCT GTATTGTATC CGATCTATAA ACTAAAATAA TTGATATAAA
AAAAAAAAAA AA

SEQ ID NO:13 NUCLEIC ACID SEQUENCE OF YL-3

AAACAACAAT ATTACAAAAA ACTATGGCAT TCAAAACGTG TGGTTTTTCA
AAAAAATGGT TAATAATAGC AGTTATAGTT ATGTGTTTGT GTACCGAGTG
TTATTGCCAA TGCACTGGGG GTGCTGATTG TACTAGTTGT ACAGCAGCAT
GCACTGGTTG TGGAAGTTGT CCAAATGCGC ATACGTGTAC CGATTCTAAA
AATTGTGTCA GGGCAGAAAC GTGTACCGAT TCTGAAAATT GTGTCAAGGC
ACATACATGT ACTGGATCTA GAAACTGTAA TACAGCCATG ACGTGTACAA
ACTCAAAAGA CTGTTTTGAA GCCAAAACAT GTACTGACTC AACCAACTGT
TACAAAGCTA CAGCCTGTAC CAATTCAACA GGATGTCCCG GACATTAAGT
TTTTCTATTG TCAACAATAA TAAAACACGG AGGGATAGTC TAAGCTAAAA
CATAATTGTA AGCTTACTCT GTATTGTATC CGATCTGTCT CTTTGCCTCC
CAAGGATGAT AATTTTGTAC TGGGAGCGAA AGGGTTACCG GACAATAATA
ATTAATAAAC TAAATAATTG ATAAAAAAAA AAAAAAA

A10

SEQ ID NO:14 NUCLEIC ACID SEQUENCE OF SEQUENCE OF YL-4

CAAAAAAGTA TGTCATTCAA AATAAGTACT TTTACAAAAA TCTGGTTAAT
TATAGCAGTT ATGTTATGTG TTTGTGTAAC GAGTATAATT GCCAGTGCAC
TGGGGCTGCT GATTGTACTA GTTGTACAGC AGCATGCACT GGTTGTGGAA
ACTGTCCAAA TGCAATAACG TGTACCGGTT CTAAAAATTG TGTCAGGGCA
ACAACATGTA CTGGGTCTAC AAACTGTAAT GAGCCACGAC GTGTACAAAT
CAAAAGGCTG TTTAGAAGCC ACAACATGTA CTGGGTCTAC ACACTGTCAT
AGAGCACGAC GTGTACAAAT TCAAAAGACT GTTTTGAAGC CACAACATGT
ACTGGCTCAA GCAACTGTTA CACTGCTACA ACATGTACTA ACTCAACCAA
CTGTTACAAA GCTACAGCCT GTACCAATTC AACAGGATGT CCCGGACATT
AGGTTTTTTT ATTGTCAACA ATAAAATAAA ACAAAACTGT TCTTATCTAA
GCTAAAACAT AAATGTAAAC GTTAATTTGT ATTCTATCCG ATCTGTCCCT
TTGCGCCCTA AGGATAATTT TGTACAGGGA GAGAAAAGGC TATCGGACAA
TAATAAACAT TGTTAATATA CATAAAAAAA AAAAAAAA

SEQ ID NO:15 NUCLEIC ACID SEQUENCE OF 5-15

AAACAGCGAG ATAAACAACA ATATTACAAA AAACT**ATG**GC ATTCAAAACG
TGTGGTTTTT CAAAAAAATG GTTAATAATA GCAGTTATAG TTATGTGTTT
GTGTACCGAG TGTTATTGCC AATGCACTGG GGGTGCTGAT TGTACTAGTT
GTACAGCAGC ATGCACTGGT TGTGGAAGTT GTCCAAATGC GCATACGTGT
ATCGATTCTA AAAATTGTGT CAGGGCAGAA ACGTGTACCG ATTCTGAAAA
TTGTGTCAAG GCACATACAT GTACTGGATC TAGAAACTGT AATACAGCCA
TGACGTGTAC AAACTCAAAA GACTGTTTTG AAGCCAAAAC ATGTACTGAC
TCAACCAACT GTTACAAAGC TACAGCCTGT ACCAATTCAA CAGGATGTCC
CGGACAT**TAA** GTTTTTCTAT TGTCAACAAT AATAAAACAC GGAGGGATAG
TCTAAGCTAA AACATAATTG TAAGCTTACT CTGTATTGTA TCCGATCTGT
CTCTTTGCCT CCCAAGGATG ATAATTTTGT ACTGGGAGCG AAAGGGTTAC
CGGACAATAA TAATTAATAA ACTAAAATAA TTGATATAAA AAAAAAAAAA
AA

SEQ ID NO:16 NUCLEIC ACID PROBE USED TO SCREEN cDNA LIBRARY

CACTGCACTG GGGGTGCTGA TTGTACTAGT TGTACAGATG CATGCACTGG
TTGTGGAAAT TGTCCAAATG CACATACGTG TACCGATTCC AAAAATTGTG
TCAAGGCAGC AACATGTACT GGATCTACAA AATGTAATAC CGCCAGACGT
GTACAAACTC AAAAGACTGT TTTGAAGCCA AAACATGTAC TGACTCAACC
AACTGTTACA AAGCTACAGC CTGTACCAAT TCAACAGGAT GTCCCGGACA
TTAAGTTTTT CTATTGTCAA CAATAATAAA ACACACTTAC TGTTATCTTA
GCTAAAACAT AA

ATG = initiating met

TAA = stop codon